

CLAIMS:

1. A nucleic acid sequencing method comprising:
providing a DNA sample containing a plurality of
5 circular single-stranded DNA template molecules each
comprising a primer annealing sequence and a target
sequence;
forming a random array of immobilized and amplified
template molecules, by
10 contacting said template molecules with an
amplification primer to anneal to the primer
annealing sequence thereby forming annealed
primer/template complexes,
amplifying said template molecules by
15 rolling-circle amplification,
ensuring said amplified template molecules
are immobilized on a solid support by immobilizing
the amplification primer before annealing the
template, the primer/template complexes before
20 amplification, or the amplified templates after
amplification;
probing the tandem-repeated amplification product
with a panel of probes under test conditions, determining
for each probe whether it hybridizes to the target sequences
25 or not under the test conditions, thereby obtaining a
hybridization spectrum of the target;
comparing the hybridization spectrum to a
hybridization spectrum for reference sequences in a
reference database comprising a plurality of reference
30 sequences, wherein the reference database is expected to
contain within it one or more reference sequences for the
sequence of the DNA template, thereby determining the likely
location or locations of the target sequence within one or
more reference sequences;

optionally computing the likely sequence of the target sequence and/or a difference in sequence of the target sequence compared with one or more reference sequences by comparing the actual hybridization spectrum
5 with the expected hybridization spectrum at the location or locations.

2. A method according to claim 1 comprising computing a difference in sequence of the target sequence compared
10 with one or more reference sequences, wherein the difference is one or more or a combination of differences selected from the group consisting of single nucleotide polymorphism, insertion, deletion, alternative splicing, an alternative transcriptional start site, alternative polyadenylation, and
15 microsatellites.

3. A method according to claim 1 or claim 2 wherein the panel of probes comprises probes with an effective specificity of 3 to 10 bases.

20

4. A method according to claim 3 wherein said effective specificity is 4 to 6 bases.

5. A method according to any one of claims 1 to 4
25 wherein the size of each target sequence and the effective specificity of the full or partial panel of probes are adjusted so that the statistical probability of hybridization of each probe to each target is between 5% and
95%.

30

6. A method according to claim 5 wherein said statistical probability is between 10% and 90%.

7. A method according to claim 6 wherein said statistical probability is between 25% and 75%.

8. A method according to claim 7 wherein said statistical probability is between 40% and 60%.

9. A method according to any one of claims 1 to 8 comprising probing with multiple panels of probes, where each probe in each panel of probes is different from each probe in each other panel of probes.

10. A method according to any one of claims 1 to 9 wherein the reference database is compiled from sequences of nucleic acid from the same species as the target sequence.

11. A method according to any one of claims 1 to 9 wherein the reference database is compiled from sequences of nucleic acid from a different species from the target sequence.

12. A method according to any one of the preceding claims comprising forming a random array of single-stranded DNA molecules, wherein

each said molecule consists of at least two tandem-repeated copies of an initial sequence,

each said molecule is immobilized on a surface at random locations with a density of between 10^3 and 10^7 per cm^2 ,

each said initial sequence represents a random fragment from an initial target DNA or RNA library comprising a mixture of single- or double-stranded RNA or DNA molecules,

said initial sequences of all said DNA molecules have approximately the same length.

13. A method according to claim 12 wherein each molecule comprises at least 1000 tandem-repeated copies of an initial sequence.
- 5
14. A method according to claim 12 or claim 13 wherein said density is between 10^5 per cm^2 and 10^7 per cm^2 .
15. A method according to any one of claims 12 to 14
- 10 wherein said initial sequences have the same length within 50% CV.
16. A method according to claim 15 wherein said initial sequences have the same length within 10% CV.
- 15
17. A method according to claim 16 wherein said initial sequences have the same length within 5% CV.
18. A method according to any one of claims 12 to 17
- 20 wherein said initial target library is an RNA library, an mRNA library, a cDNA library, a genomic DNA library, a plasmid DNA library or a library of DNA molecules.
19. A method according to any one of the preceding
- 25 claims wherein, in the panel of probes:
- each probe consists of one or more oligonucleotides,
 - each said oligonucleotide is stabilized,
 - each said oligonucleotide carries a reporter moiety,
 - the effective specificity of each probe is between 3
- 30 and 10 bp,
- the set of probes is such that at least 10% of all positions in a random or arbitrary target sequence statistically hybridize with at least one probe in the set of probes.

20. A method according to claim 19 wherein the effective specificity is between 4 and 6 bp.
- 5 21. A method according to claim 19 or claim 20 wherein the panel of probes statistically hybridizes to at least 25% of all positions in a target sequence.
22. A method according to claim 21 wherein the panel of
10 probes statistically hybridizes to at least 50% of all positions in a target sequence.
23. A method according to claim 22 wherein the panel of probes statistically hybridizes to at least 90% of all
15 positions in a target sequence.
24. A method according to claim 23 wherein the panel of probes statistically hybridizes to 100% of all positions in a target sequence.
- 20 25. A method according to any one of claims 19 to 24 stabilised by one or more of introduction of degenerate positions, introduction of locked nucleic acid monomers, introduction of peptide nucleic acid monomers and
25 introduction of a minor groove binder.
26. A method according to any one of claims 19 to 25 wherein the reporter moiety is selected from the group consisting of a fluorophor, a quencher, a dark quencher, a
30 redox label, and a chemically reactive group which can be labeled by enzymatic or chemical means, for example a free 3'-OH for primer extension with labeled nucleotides or an amine for chemical labelling after hybridization.

27. A method according to any one of the preceding claims, wherein the hybridisation spectra are compared using a spectral search instrument comprising a field-programmable gate array (FPGA) attached to a host computer and a
5 computer-readable memory device, wherein
said FPGA is configured to perform spectral search,
said computer-readable memory device stores a reference nucleotide sequence and a set of hybridization spectra,
said host computer is configured to provide said FPGA
10 with the reference nucleotide sequence and with each said hybridization spectrum,
said FPGA, when provided with a reference nucleotide sequence and a hybridization spectrum, writes to said computer-readable memory to store the location or locations
15 of best matches between said hybridization spectrum and said reference nucleotide sequence.

28. A computer processor programmed to control a method of according to any one of claims 1 to 27.

20

29. A computer-readable device carrying a program for a computer processor according to claim 28.

30. A computer processor programmed to provide sequence
25 information for a nucleic acid from performance of a method according to any one of claims 1 to 27.

31. A computer-readable device carrying a program for a computer processor according to claim 30.

30

32. A random array of single-stranded DNA molecules,
wherein

each said molecule consists of at least two tandem-repeated copies of an initial sequence,

each said molecule is immobilized on a surface at random locations with a density of between 10^3 and 10^7 per cm^2 ,

each said initial sequence represents a random
5 fragment from an initial target DNA or RNA library comprising a mixture of single- or double-stranded RNA or DNA molecules,

said initial sequences of all said DNA molecules have approximately the same length.

10

33. A random array according to claim 32 wherein each molecule comprises at least 1000 tandem-repeated copies of an initial sequence.

15 34. A random array according to claim 32 or claim 33 wherein said density is between 10^5 per cm^2 and 10^7 per cm^2 .

35. A random array according to any one of claims 32 to 34 wherein said initial sequences have the same length
20 within 50% CV.

36. A random array according to claim 35 wherein said initial sequences have the same length within 10% CV.

25 37. A random array according to claim 36 wherein said initial sequences have the same length within 5% CV.

38. A random array according to any one of claims 32 to 37 wherein said initial target library is an RNA library, an
30 mRNA library, a cDNA library, a genomic DNA library, a plasmid DNA library or a library of DNA molecules.

39. A set of probes wherein each probe consists of one or more oligonucleotides,

each said oligonucleotide is stabilized,
each said oligonucleotide carries a reporter moiety,
the effective specificity of each probe is between 3
and 10 bp,

5 the set of probes is such that at least 10% of all
positions in a random or arbitrary target sequence
statistically hybridize with at least one probe in the set
of probes.

10 40. A set of probes according to claim 39 wherein the
effective specificity is between 4 and 6 bp.

41. A set of probes according to claim 39 or claim 40
which statistically hybridizes to at least 25%, at least
15 50%, at least 90% of all positions in a target sequence.

42. A set of probes according to claim 41 which
statistically hybridizes to 100% of all positions in a
target sequence.

20

43. A set of probes according to any one of claims 39
to 42 stabilised by one or more of introduction of
degenerate positions, introduction of locked nucleic acid
monomers, introduction of peptide nucleic acid monomers and
25 introduction of a minor groove binder.

44. A set of probes according to any one of claims 39
to 43 wherein the reporter moiety is selected from the group
consisting of a fluorophor, a quencher, a dark quencher, a
30 redox label, and a chemically reactive group which can be
labeled by enzymatic or chemical means, for example a free
3'-OH for primer extension with labeled nucleotides or an
amine for chemical labelling after hybridization.

45. A spectral search instrument comprising a field-programmable gate array (FPGA) attached to a host computer and a computer-readable memory device, wherein

said FPGA is configured to perform spectral search,

5 said computer-readable memory device stores a reference nucleotide sequence and a set of hybridization spectra,

said host computer is configured to provide said FPGA with the reference nucleotide sequence and with each said hybridization spectrum,

10 said FPGA, when provided with a reference nucleotide sequence and a hybridization spectrum, writes to said computer-readable memory to store the location or locations of best matches between said hybridization spectrum and said reference nucleotide sequence.

15